

Monthly Consumers' Complaints reported to Consumer Financial Protection Bureau: Utilization of Automatic Keywords Extraction Techniques

Wookjae Heo, South Dakota State University¹

Narang Park, University of Georgia²

Lorna Saboe-Wounded Head, South Dakota State University³

Objective and Significance

According to Consumer Financial Protection Bureau (CFPB), over 500,000 complaints were received from 2015 – 2017. Thirty-five percent of the complaints included narratives that provided valuable qualitative data to aid in analyzing the complaints. With advanced technology from the field of artificial intelligence computations, it can be possible to examine the large qualitative database to better understand complaint themes. A mixed methodology (i.e., text analysis) can be used to analyze the complaint database. As a result, the objective of this study is to determine the most used monthly keywords by utilizing text analytic techniques to bring significant data for researchers, Cooperative Extension educators, and policy makers to understand the seasonal financial issues consumers face.

Method

This study analyzed consumer complaint narratives submitted to the CFPB website from 2015 - 2017. Consumers' complaints have been recorded after the company responded to the case or 15 days later (CFPB, 2018). In 2015, 152,977 complaints were recorded, 39,239 consumers included narratives with their complaint. In 2016, 169,342 complaints were submitted with 55,688 narratives. Two thousand seventeen yealed 210,128 complaints with 92,326 narratives. As a result, 187,253 consumer narratives were analyzed using the text analytic technique.

The text analytic method uses both qualitative and quantitative data (Bauer, Bicquelet, & Serdem, 2014). From the qualitative database of narratives, it is possible to find meaningful combinations of key words by utilizing quantitative calculations (Roberts, 1997). Then a semantic technique is used to identify most used keywords. This technique is called Automatic Keyword Extraction (AKE), introduced by Rose et al. (2010). The basic algorithms of AKE are based on findings manifested by words in a document, which calculates the most dominating words in a same phrase or in a same document. The text analytic technique is widely used in diverse fields. For example, it is possible to find psychological traits from a personal narrative (Ignatow & Mihalcea, 2018; Mairesse et al., 2007; Pennebaker & Kin, 1999) and to examine social relationships by analyzing conversations (Stark, Shafran, & Kaye, 2012). For this study, all narratives are classified into monthly database and analyzed for finding seasonal trend of consumer issues in the

¹ Wookjae Heo, (wookjaeheo@gmail.com), Assistant Professor, Department of Consumer Sciences

² Narang Park, (narang.park25@uga.edu), Ph.D student, Department of Financial Planning, Housing and Consumer Economics

³ Lorna Saboe-Wounded Head, (lorna.woundedhead@sdstate.edu), Family Resource Management Field Specialist, Department of Consumer Sciences

consumer financial market.

Findings

Table 1 shows the most frequent nouns in each month's narrative. Annually, there are similar topics among consumer complaints. However, thinking that financial consumers are struggling with credit score, debt, loan, and payment(s), the top 10 nouns in all months are commonly acceptable as general issues.

However, as showing on Table 2, the monthly key words found by AKE algorithms are more meaningful than the top 10 nouns. Specifically, there has been definitely seasonal trend in consumers' complaints. Even though the Table 2 shows only total result of dominating key words from 3-years' narratives, the dominating monthly key words in each year (i.e., monthly key words in 2015, in 2016, and in 2017) showed similar pattern. The detailed result will be shown on the conference presentation.

The analyses reveal that the keywords regarding debt collection (e.g., 1692g sec, Act allow, financial aid), mortgage (e.g., loss mitigation, adjustable rate), and credit reporting (e.g., social security, major credit, free monitoring, credit worthy, credit scoring) are generally and consistently shown across the months. This result is supported by the CFPB's consumer response annual report (2018) that debt collection, mortgage, and credit reporting are the top three financial products or services submitted in the consumer complaints database over the past three years from 2015 to 2017.

There are some keywords found relevant to the specific season or issues. For example, "Tax preparer" is the keyword that appears in March. As many people prepare their tax reports before the due date in April, their complaints about tax preparers tend to increase in March. "Air conditioner" is one of the frequently found keywords in April. This is plausible because consumers are likely to buy air conditioners to get ready for the summer. Also, a keyword like "International Student" uniquely appears in September. Most of the international students make flights to the United States in August or September to begin their college, and new comers are likely to open bank accounts with various promotions. International students are likely to report complaints regarding unsatisfactory financial services from banks in the following month. Issues that occurred at that time of the year are reflected in the dominating monthly keywords.

The noteworthy keywords found over the three years are cyber security, digital security, and lax security. As many transactions are digitalized and personal information is transferred via online/mobile, digital security is increasingly gaining attention. The reason that the relevant keywords are mainly found in September is that there was a cyberattack on Equifax in 2017. According to the NYT (Sept 7, 2017), at least 143 million Americans' data, including social security numbers and driver's license numbers, was leaked from this attack. The complaints filed by the furious consumers in that year seem to contribute to finding dominant keywords from the database. Due to this issue, "Privacy act" is found as one of the dominating keywords in October.

Conclusion/Relevance

The Consumer Finance Protection Bureau provides a valuable resoruces for consumers.

The ability to file a complaint and to potentially receive help in dealing with the issue is benefit for consumers. Utilizing the text analytic method to identify predominate consumer issues can inform researchers and educators about emerging issues so resources and solutions can be developed.

Table 1. Monthly top 10 nouns manifest in consumer complaints narratives from 2015 to 2017.

<u>January</u>		<u>February</u>		<u>March</u>		<u>April</u>		<u>May</u>		<u>June</u>	
Noun	Freq.	Noun	Freq.	Noun	Freq.	Noun	Freq.	Noun	Freq.	Noun	Freq.
Payment(s)	12,734	Payment(s)	10,211	Credit	11,793	Credit	16,841	Credit	13,343	Credit	19,136
Credit	11,085	Credit	10,068	Account	10,636	Payment(s)	15,206	Account	12,452	Account	17,895
Account	10,253	Account	9,619	Payment(s)	10,539	Account	14,857	Payment(s)	10,542	Payment(s)	15,504
Loan	8,292	Loan	6,497	Loan	6,556	Loan	10,194	Loan	7,240	Information	9,265
Time	4,894	Debt	4,365	Debt	5,441	Information	7,490	Debt	5,390	Debt	7,609
Information	4,873	Information	4,204	Information	5,216	Debt	7,407	Information	5,269	Report	7,605
Debt	4,817	Time	4,023	Report	4,633	Report	6,810	Report	5,064	Time	6,867
Company	4,129	Company	3,681	Time	4,332	Time	6,270	Time	4,933	Company	6,563
Report	3,891	Report	3,628	Company	4,174	Company	5,709	Company	4,513	Card	5,710
<u>July</u>		<u>August</u>		<u>September</u>		<u>October</u>		<u>November</u>		<u>December</u>	
Noun	Freq.	Noun	Freq.	Noun	Freq.	Noun	Freq.	Noun	Freq.	Noun	Freq.
Credit	6,319	Credit	22,253	Credit	24,508	Credit	21,885	Credit	19,198	Credit	18,791
Account	5,800	Account	19,892	Account	19,785	Account	19,488	Account	17,056	Account	17,863
Payment(s)	5,403	Payment(s)	17,521	Payment(s)	16,845	Payment(s)	16,714	Payment(s)	14,851	Payment(s)	15,106
Loan	3,563	Loan	10,759	Information	13,601	Loan	9,817	Loan	8,511	Loan	8,517
Debt	2,819	Information	9,291	Loan	10,594	Information	8,916	Information	7,945	Information	8,390
Information	2,535	Debt	9,070	Report	8,385	Report	8,631	Report	7,250	Debt	7,529
Time	2,511	Report	9,053	Debt	7,836	Debt	7,868	Debt	6,757	Report	7,445
Report	2,398	Time	7,390	Time	7,632	Time	7,163	Company	6,212	Company	6,659
Card	2,138	Company	7,369	Company	7,433	Company	7,032	Time	6,156	Card	6,161
Company	2,077	Letter	6,471	Card	6,931	Card	6,626	Card	5,831	Time	6,124

Table 2. Monthly dominating key words found by AKE algorithms from consumer complaints narratives from 2015 to 2017.

January	February	March	April	May	June
Social security	Social security	Social security	Act allow	Navy federal	Account umber
Act allow	Act allow	Act allow	Wet link	Social security	Social security
Real estate	Cellular service	Western union	Social security	Real estate	Company
Monitoring service	Western union	Tax preparer	Company	Preservation team	Payment
American express	Loss mitigation	Navy federal	National recovery	Lump sum	Western union
Credit monitoring	Financial security	GC services	Payment	Portfolio recovery	Social media
Loss mitigation	Best buy	Portfolio recovery	Adjustable rate	Social media	Installment account
1692g sec.	Real estate	Act guideline	Air conditioner	Monitoring service	Customer relations
3 rd party	Act	Vice president	Financial security	Vice president	Original insert
Tran union	Cellular telephone	Loss mitigation	Private student	American express	Zip code
July	August	September	October	November	December
Social security	Social security	Act allow	Act allow	Act allow	Act allow
Navy federal	Act laws	Social security	Privacy act	Social security	Credit worthy
Green tree	Navy federal	Cyber security	Social security	Navy federal	Social security
Western union	GC services	Free monitoring	Navy federal	CRA violation	American express
Management service	Adjustable rate	Digital security	Great lakes	Western union	Loss mitigation
Major credit	Credit monitoring	Prepaid card	Vice president	Social media	Recovery solutions
Loss mitigation	Mitigation plan	Lax security	Financial aid	Credit scoring	5 th 3 rd
Real estate	Financial security	Real estate	Customer relations	Real estate	Act law
Best buy	Social service	Square footage	Real estate	Hall runner	Secret key
American express	Loss mitigation	International student	American express	Social number	Great lakes

References

- Bauer, M. W., Bicuquelet, A., & Suerdem, A. K. (Eds.). (2014). Text analysis: An introductory manifesto. In < W. Bauer, A. Bucquelet, & A. K., Suerdem (Eds.), *Textual analysis (SAGE benchmarks in social research methods)* (Vol. 1). Thousand Oaks, CA: SAGE publications.
- Consumer Financial Protection Bureau. (2018). Consumer Complaint Database. Retrieve from <https://www.consumerfinance.gov/data-research/consumer-complaints/>
- Ignatow, G., & Mihalcea, R. (2018). *An Introduction to Text Mining: Research Design, Data Collection, and Analysis*. Thousand Oaks: SAGE publications.
- Mairesse, F., Walker, M., Mehl, M., & Moore, R. (2007). Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research*, 30, 457-501.
- Pennebaker, J. W., & King, L. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, 77, 1296-1312.
- Roberts, C. W. (1997). *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts*. Mahwah, NJ: Lawrence Erlbaum.
- Rose, S., Engel, D., Cramer, N., & Cowley, W. (2010). Automatic Keyword Extraction from Individual Documents. In M. W. Berry & J. Kogan, *Text Mining: Applications and Theory* (pp. 3-20). John Wiley & Sons, Ltd.
- Stark, A., Shafran, I., & Kaye, J. (2012). Hello, who is calling?: Can words reveal the social nature of conversations? *Proceedings of the 2012 Conference of the North American Chapters of the Association for Computational Linguistics: Human Language Technologies* (pp. 112-119).
- Consumer Financial Protection Bureau (2018) Consumer response annual report. Retrieved from https://s3.amazonaws.com/files.consumerfinance.gov/f/documents/cfpb_consumer-response-annual-report_2017.pdf
- Bernard, T. S., Hsu, T., Perlroth, N., & Lieber, R. (September 2017). Equifax says cyberattack may have affected 143 million in the U.S. *The New York Times*. Retrieved from <https://www.nytimes.com/2017/09/07/business/equifax-cyberattack.html>